ELSEVIER

Theoretical Computer Science 233 (2000) 91–106

Theoretical Computer Science

www.elsevier.com/locate/tcs

polyatomic structures by X-rays

R.J. Gardner [a,*,1], P. Gritzmann [b,2], D. Prangenberg [c,3]

[a] *Western Washington University, Department of Mathematics, Bellingham, WA 98225-9063, USA*
[b] *Technische Universität München, Zentrum Mathematik, D-80290 München, Germany*
[c] *Universität Trier, Fachbereich IV, Mathematik, Postfach 3825, D-54286 Trier, Germany*

Received May 1997; revised December 1997
Communicated by M. Nivat

Abstract

The problem of recovering the structure of crystalline materials from their discrete X-rays is of fundamental interest in many practical applications. An important special case concerns determining the position of atoms of several different types in the integer lattice, given the number of each type lying on each line parallel to some lattice directions. We show that the corresponding consistency problem is $\mathbb{NP}$-complete for any two (or more) different (fixed) directions when six (or more) types of atoms are involved.

AMS classification: 05; 68R; 69U; 82D; 90C; 92C

Keywords: Tomography; Discrete tomography; Computational complexity; Polynomial-time algorithm; $\mathbb{NP}$-completeness; Lattice; Contingency table; Data security; Scheduling

1. Introduction

Motivated by applications in biology, medicine and the material sciences, the area of tomography deals with recovering the structure of objects from their X-rays. The most familiar of these applications is the CAT scanner, an invaluable tool in medicine for diagnostic and surgical purposes. The mathematics of computerized tomography is well developed and enables the processing of the X-ray information to produce approximate sectional images.

* Corresponding author. E-mail: gardner@baker.math.wwu.edu.

[1] Supported in part by the Alexander von Humboldt Foundation and by National Science Foundation Grant DMS-9501289.

[2] Supported in part by the Deutsche Forschungsgemeinschaft and by a Max Planck Research Award.

[3] Supported by the Deutsche Forschungsgemeinschaft through the "Graduiertenkolleg Mathematische Optimierung" at the Universität Trier.

0304-3975/00/$ - see front matter
PII: S0304-3975(97)00298-3

A different branch of tomography called discrete tomography is rapidly evolving in the attempt to find suitable algorithms to process discrete X-ray data. Such algorithms will enjoy immediate application in the material sciences, for example in the problem of determining molecular structures. Recent improvements in High Resolution Transmission Electron Microscopy (HRTEM) has led to the generation of data with resolutions on the atomic scale. In particular, the papers [11, 14] describe a new technique based on HRTEM that can effectively measure the number of atoms in a crystal lying on each line parallel to certain directions. This technique is called *QUANTITEM* (QUantitative ANalysis of The Information from Transmission Electron Microscopy). We call such a measurement a discrete X-ray, since it is indeed a discrete form of the conventional X-ray. (See Section 2 for precise definitions.) The goal is to use QUANTITEM to reconstruct crystalline structures, and this leads to the mathematical inverse problem of reconstructing finite sets of points in a lattice from their discrete X-rays.

When there is only one type of atom present in the crystal, the problem has already been studied quite extensively. In particular, its computational complexity has been determined completely in [6]. Briefly, the problem of reconstructing a finite subset of the integer lattice $\mathbb{Z}^d$ ($d \geqslant 2$) from discrete X-rays in $m \geqslant 2$ arbitrary but fixed pairwise nonparallel lattice directions is solvable in polynomial time if $m = 2$, and is $\mathbb{NP}$-hard if $m \geqslant 3$. The restriction to lattice directions is not an artificial one, since the QUANTITEM technique can only provide discrete X-rays in certain lattice directions.

The present paper studies the complexity of reconstructing materials composed of $c \geqslant 2$ types of atoms from discrete X-rays in $m \geqslant 2$ pairwise nonparallel lattice directions. (The discrete X-rays now give the number of atoms of each type lying on each line parallel to the directions.) We refer to this as the *polyatomic* case. This problem is of fundamental importance in practice, for example in determining interfaces in materials. By the known results for the homogeneous case $c = 1$, this problem is $\mathbb{NP}$-hard for $m \geqslant 3$, so it suffices to restrict all considerations to the case $m = 2$.

Our main theorems, Theorems 3.2 and 3.5, show that the corresponding consistency and uniqueness problems are $\mathbb{NP}$-complete if $c \geqslant 6$. The final section of the paper contains further remarks, including consequences for some problems in statistics, data security, and machine scheduling.

2. Definitions, main problems, and preliminaries

For a set A, we denote by $\operatorname{lin} A$ the *linear span* of A. The symbol $\mathbb{1}_A$ represents the *characteristic function* of A.

For $d, k \in \mathbb{N}$ with $d \geqslant 2$ and $k \leq d-1$, let $\mathscr{S}_{k,d}$ be the set of all k-dimensional subspaces in Euclidean d-space $\mathbb{E}^d$. If $S \in \mathscr{S}_{k,d}$, $S^\perp$ denotes the orthogonal complement of S. (Though we only require the special case $k = 1$ here, this notation conforms to that of some other papers in this area.)

Let F be a finite subset of $\mathbb{E}^d$, let $S \in \mathscr{S}_{1,d}$, and let $\mathscr{A}(S)$ denote the set of all lines parallel to S. The (*discrete*) *X-ray of F parallel to S* is the function $X_S F : \mathscr{A}(S) \to$

$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ defined by

$$X_S F(T) = |F \cap T| = \sum_{x \in T} \mathbb{1}_F(x),$$

for $T \in \mathscr{A}(S)$.

Since we only consider discrete X-rays in this paper, we shall simply refer to them as X-rays. Each X-ray provides line sums that count the number of elements of a finite set on each line parallel to the given 1-dimensional subspace. It is in effect the projection, counted with multiplicity, of F on $S^{\perp}$.

Most of our results extend to arbitrary finite subsets of $\mathbb{E}^d$, but we shall focus on finite subsets of a given rational lattice, a subset of $\mathbb{E}^d$ that consists of all integer combinations of a fixed set of d linearly independent rational vectors. If the lattice is known explicitly, as is the case for the crystalline structures that are to be determined by QUANTITEM, the affine invariance of the problems we consider allows us to assume that the lattice is the integer lattice $\mathbb{Z}^d$. Therefore we shall work exclusively in $\mathbb{Z}^d$ in the sequel.

Let $\mathscr{L}_{1,d}$ denote the subset of $\mathscr{S}_{1,d}$ consisting of lines spanned by a nonzero vector of $\mathbb{Z}^d$, and let $\mathscr{F}^d$ denote the class of all finite subsets of $\mathbb{Z}^d$. Each element of $\mathscr{L}_{1,d}$ or $\mathscr{F}^d$ will be called a *lattice line* or *lattice set*, respectively.

In the sequel we consider only X-rays, parallel to a 1-dimensional lattice subspace, of lattice sets. Each such X-ray vanishes except on a finite family of parallel lattice lines.

Throughout the paper, we shall assume that $d, m, c \in \mathbb{N}$ with $d, m \geqslant 2$, and that $S_1, \ldots, S_m \in \mathscr{L}_{1,d}$ are different. We shall also use the notation

$$S_i^* = \operatorname{lin}\{e_i\} \quad \text{for } i = 1, 2, \text{ where } e_1 = (1,0), e_2 = (0,1),$$

for the coordinate lines in the plane.

We now state the first algorithmic problem.

POLY$_c$-CONSISTENCY$_{\mathscr{F}^d}(S_1, \ldots, S_m)$

Instance: *For each* $l = 1, \ldots, c$ *and* $i = 1, \ldots, m$, *a function* $f_{i,l} : \mathscr{A}(S_i) \to \mathbb{N}_0$ *with finite support.*

Question: *Do there exist disjoint sets* $F_l \in \mathscr{F}^d$ *such that* $X_{S_i} F_l(T) = f_{i,l}(T)$ *for* $l = 1, \ldots, c$ *and* $i = 1, \ldots, m$*?*

The statement of POLY$_c$-CONSISTENCY$_{\mathscr{F}^d}(S_1, \ldots, S_m)$ is slightly ambiguous since we have not specified the data structures for encoding $f_{i,l}$ precisely. This can be done by encoding each such function as a finite subset of $\mathbb{Z}^d \times \mathbb{N}$ of minimal cardinality. The paper [6] provides a detailed rigorous algorithmic description of various aspects of this data structure (for $c = 1$). We omit the details here, but note that with the appropriate data structure in place, it is clear that POLY$_c$-CONSISTENCY$_{\mathscr{F}^d}(S_1, \ldots, S_m)$ is in the class $\mathbb{NP}$.

In previous papers [6, 7] we considered the case $c = 1$, which corresponds to a homogeneous material. Here we are interested in the inhomogeneous case $c \geqslant 2$. For each $l = 1, \ldots, c$ and $i = 1, \ldots, m$, the instance provides a candidate $f_{i,l}$ for the X-ray parallel to S_i of an unknown lattice set F_l whose points represent atoms of the lth type. Clearly, the sets F_l must be disjoint to model physical reality.

Note that a necessary condition for the consistency of the given instance is that for each l, all sums $\sum\{f_{i,l}(T) : T \in \mathscr{A}(S_i)\}$ are equal for each i to some $N_l \in \mathbb{N}$, the cardinality of any solution F_l.

The reconstruction problem Poly$_c$-Reconstruction$_{\mathscr{F}^d}(S_1, \ldots, S_m)$ is defined in a way similar to Poly$_c$-Consistency$_{\mathscr{F}^d}(S_1, \ldots, S_m)$. The input is the same but the question is replaced by the task of constructing a solution. We also consider the following algorithmic task related to uniqueness.

Poly$_c$-Uniqueness$_{\mathscr{F}^d}(S_1, \ldots, S_m)$

Instance: *Disjoint* $F_l \in \mathscr{F}^d$, *where* $l = 1, \ldots, c$.

Question: *Do there exist disjoint* $F'_l \in \mathscr{F}^d$ *such that* $X_{S_i}F_l = X_{S_i}F'_l$ *for* $l = 1, \ldots, c$ *and* $i = 1, \ldots, m$, *and* $F'_{l_0} \neq F_{l_0}$ *for some* l_0?

Note that Poly$_c$-Uniqueness$_{\mathscr{F}^d}(S_1, \ldots, S_m)$ is in the class $\mathbb{NP}$.

When $c = 1$ and $m \geqslant 3$, the consistency and uniqueness problems are $\mathbb{NP}$-complete and the reconstruction problem is $\mathbb{NP}$-hard; see [6]. We conclude immediately that the same is true for $c \geqslant 2$. When $c = 1$ and $m = 2$, each of these three problems can be solved in polynomial time; see [2, 7, 8, 13]. In view of this we assume in the sequel that $m = 2$ and $c \geqslant 2$. We remark that the earlier paper [3] also considers the case $c = 2$, but presents no complexity results.

3. Main complexity results

Our first aim is to prove that Poly$_c$-Consistency$_{\mathscr{F}^d}(S_1, S_2)$ is $\mathbb{NP}$-complete for any subset $\{S_1, S_2\}$ of $\mathscr{L}_{1,d}$ and any $c \geqslant 6$. The first step is to reduce this task to the case of the coordinate lines S_1^*, S_2^* in the plane. The following lemma is similar to [6, Lemma 3.2].

Lemma 3.1. *Suppose that* $S_1, S_2 \in \mathscr{L}_{1,d}$. *There is a polynomial-time parsimonious transformation from* Poly$_c$-Consistency$_{\mathscr{F}^2}(S_1^*, S_2^*)$ *to* Poly$_c$-Consistency$_{\mathscr{F}^d}(S_1, S_2)$. *If the former problem is* $\mathbb{NP}$*-hard in the strong sense, then so is the latter.*

Proof. Recall that S_1 and S_2 are different. Let $S_i = \operatorname{lin}\{u_i\}$, where $u_i \in \mathbb{Z}^d \setminus \{0\}$, $i = 1, 2$, and let $A : \mathbb{E}^2 \to \mathbb{E}^d$ denote the linear map defined by $Ae_i = u_i$, $i = 1, 2$. Note that A is a rank 2 matrix, so there is a matrix $B : A(\mathbb{E}^2) \to \mathbb{E}^2$ such that BA is the identity on $\mathbb{E}^2$ and AB is the identity on $A(\mathbb{E}^2)$.

Suppose that $\mathscr{I}^*$ is an instance of Poly$_c$-Consistency$_{\mathscr{F}^2}(S_1^*, S_2^*)$, and let $Z_{i,l}^* \subset \mathbb{Z}^2$ be minimal such that $Z_{i,l}^* + S_i^*$ is the support of the corresponding input function $f_{i,l}^*$ for $i = 1,2$ and $l = 1,\ldots,c$. We define an instance $\mathscr{I}$ of Poly$_c$-Consistency$_{\mathscr{F}^d}(S_1, S_2)$ by setting

$$f_{i,l}(Az + S_i) = f_{i,l}^*(z + S_i^*)$$

for $z \in Z_{i,l}^*$. Let $(F_1^*,\ldots,F_c^*)$ be a solution for $\mathscr{I}^*$ and let $F_l = AF_l^*$ for $l = 1,\ldots,c$. Then the sets F_l are pairwise disjoint, and

$$F_l \cap (Az + S_i) = A\big(F_l^* \cap (z + S_i^*)\big),$$

so $(F_1,\ldots,F_c)$ is a solution for $\mathscr{I}$.

Similarly, if $(F_1,\ldots,F_c)$ is a solution for $\mathscr{I}$, then $(BF_1,\ldots,BF_c)$ is a solution for $\mathscr{I}^*$.

Finally, note that the transformation defined above runs in strongly polynomial time and is parsimonious. □

Our main proof will describe a suitable transformation from the following variant of the well-known problem Exact Cover (see [10, p. 95]).

k-Exact Cover

Instance: *An integer k, a finite set P, a family $\mathscr{P}$ of subsets of P.*

Question: *Is there a subfamily $\mathscr{Q}$ of $\mathscr{P}$ with cardinality k such that the sets of $\mathscr{Q}$ are disjoint and the union is equal to P?*

It is shown in [10] that k-Exact Cover is $\mathbb{NP}$-complete, and [5, 15] give a parsimonious transformation from Satisfiability showing that the corresponding counting problem #(k-Exact Cover) is #$\mathbb{P}$-complete.

Theorem 3.2. Poly$_c$-Consistency$_{\mathscr{F}^2}(S_1^*, S_2^*)$ *is* $\mathbb{NP}$*-complete* (*in the strong sense*) *for* $c \geqslant 6$.

Proof. Let $\mathscr{I} = (k, P, \mathscr{P})$ be an instance of k-Exact Cover with $P = \{p_1,\ldots,p_t\}$ and $\mathscr{P} = \{P_1,\ldots,P_p\}$. We shall construct an instance $\mathscr{J}$ of Poly$_6$-Consistency$_{\mathscr{F}^2}(S_1^*, S_2^*)$ such that there is a solution for $\mathscr{J}$ if and only if there is one for $\mathscr{I}$.

Each set F_l in a solution for $\mathscr{J}$ consists of points representing atoms of a fixed type. To ease visual interpretation of the proof we assign each type of atom a color and work with lattice squares rather than lattice points. Each $z = (z_1, z_2) \in \mathbb{Z}^2$ is the center of the lattice square

$$C_z = \{x = (x_1, x_2) \in \mathbb{R}^2 : |x_j - z_j| \leqslant 1/2, j = 1,2\}.$$

A lattice square whose center represents an atom of a particular color will be assigned that color. For the six values of l corresponding to the six types of atoms, we use the

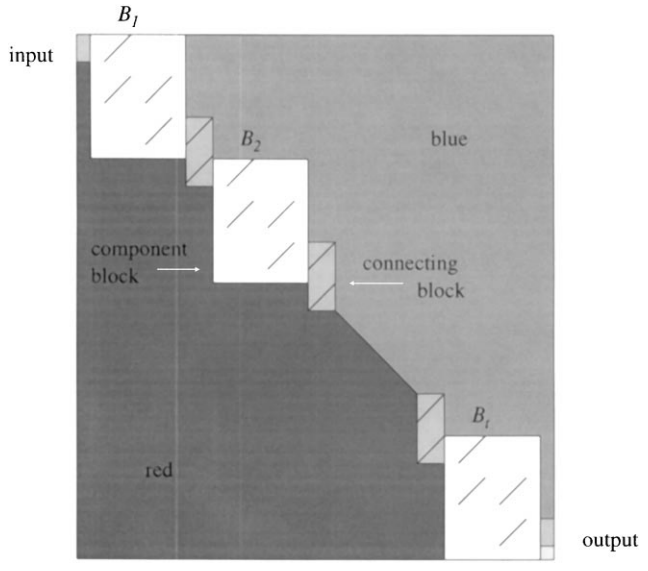

Fig. 1. The master block.

colors red, blue, green, sienna, indigo, and orange, and indicate these in the figures below by their initial letters r, b, g, s, i, and o (chosen to minimize confusion with other notation). Of course, a lattice square may be left blank to indicate that there is no atom of any color at its center.

We shall describe the instance $\mathscr{J}$ by defining six functions on rows of lattice squares specifying row sums and six functions on columns specifying column sums. It will be convenient to assign colors to the functions according to the color of the atoms they count in any solution for $\mathscr{J}$, if one exists. For example, the vertical sienna function gives column sums for sienna atoms in a solution. The reader may find the figures helpful; these illustrate the encoding of an example in which the instance $\mathscr{I}$ of k-EXACT COVER has $k = 3$, $t = 5$, $p = 7$, and

$$\mathscr{P} = \{\{p_1\}, \{p_3\}, \{p_1, p_2\}, \{p_1, p_3\}, \{p_1, p_4\}, \{p_2, p_5\}, \{p_2, p_3, p_5\}\}.$$

Note that there is a solution for this particular example, namely

$$\mathscr{Q} = \{\{p_3\}, \{p_1, p_4\}, \{p_2, p_5\}\}.$$

The functions all have value 0 on rows or columns not meeting the large rectangular block of lattice squares depicted in Fig. 1. We call this the *master block*.

The position of the master block is not important; the origin could be at its top left corner, for example. Rows are counted from the top of this (or any other) block, and columns from the left of it. The master block contains smaller rectangular blocks $B_1, \ldots, B_t$ (white in Fig. 1) called *component blocks*, each of size $(2p+3) \times (3p+3)$, and $(t-1)$ smaller *connecting blocks* (lightly shaded) of size $p \times (2p+1)$ between

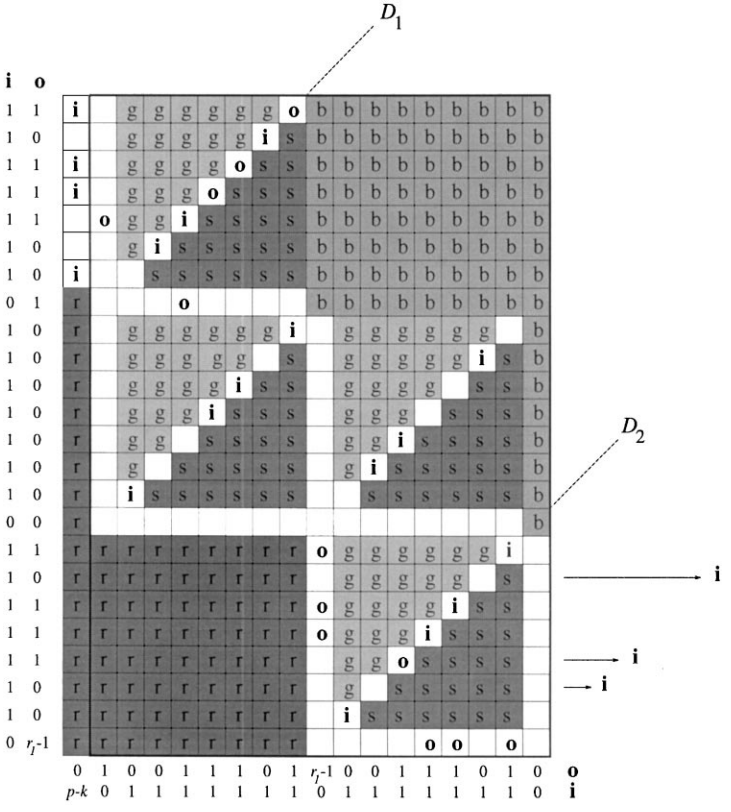

Fig. 2. The input and first component block.

them. The first component block B_1 and first connecting block are shown in Figs. 2 and 3, respectively. The left column shown in Fig. 2 is not contained in B_1. This column can be regarded as an 'input' to the master block, just as a similar exceptional vertical column of lattice squares can be regarded as the 'output'. These are indicated at the top left and bottom right of Fig. 1.

The general idea of the proof is as follows. We shall specify values for the red, blue, green, and sienna functions in such a way that there are unique sets of atoms of these colors whose X-rays agree with these functions. These sets will be such that in any solution for $\mathscr{J}$, atoms of the remaining two colors, indigo and orange, can only occupy a few diagonal, horizontal, and vertical segments (i.e., consecutive lattice squares) within the master block. (The diagonal segments play a particularly important role and their position is indicated in Fig. 1.) The component block B_i represents the element p_i of P, $i = 1,\ldots,t$. The p lattice squares in the input or on any of the special diagonal segments represent the elements of $\mathscr{P}$ in some fixed order. In the example depicted, the elements $\{p_1\},\{p_3\},\{p_1,p_2\},\{p_1,p_3\},\{p_1,p_4\},\{p_2,p_5\},\{p_2,p_3,p_5\}$ correspond to the squares in the input, counted, as always, from the top. On rows and columns

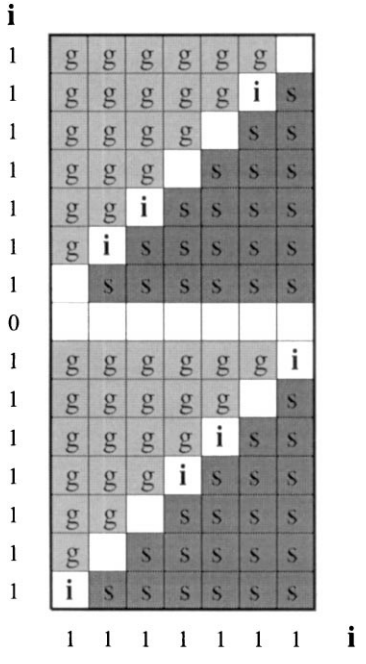

Fig. 3. The first connecting block.

meeting B_i, the values of the orange functions encode those elements of $\mathscr{P}$ that contain p_i. Thus the values of the orange function in Fig. 2 indicate that p_1 belongs to the first, third, fourth, and fifth elements of $\mathscr{P}$.

The values of the orange and indigo functions together encode elements in $\mathscr{P}$ belonging to a solution for the instance $\mathscr{I}$ of k-Exact Cover. The construction will ensure that in any particular solution for the instance $\mathscr{J}$, the relative positions of the red, blue, green, and sienna atoms is the same within each component block and also within each connecting block. Also, the relative position of the k indigo atoms on a diagonal segment meeting rows 1 through p and columns 2 through $(p+1)$ of each component block B_i (e.g., D_1 in Fig. 2) is the same and is completely determined by the $(p-k)$ indigo atoms in the input. Any solution for $\mathscr{J}$ will contain in each B_i precisely one orange atom on the same row as one of these k indigo atoms. This will mean that only the element in $\mathscr{P}$ corresponding to this indigo atom can contain $p_i \in P$. Consequently, the positions of the k indigo atoms on any of these diagonals yield a subset of $\mathscr{P}$ of cardinality k that is a solution for the instance $\mathscr{I}$ of k-Exact Cover. The converse works in a similar fashion.

We now proceed with the details of this construction. The values of the red and blue functions are chosen so that the red and blue regions in the master block (Fig. 1) represent sets whose X-rays agree with these functions. The crucial point is that these sets are unique. One way to see this is to recall that by [4, Theorem 3] a planar lattice set is uniquely determined by its X-rays parallel to the coordinate lines if and only if

it has no bad rectangle. A *bad rectangle* for a lattice set F is a set consisting of four lattice points forming a rectangle, such that the two points on one diagonal belong to F and the two on the other diagonal do not. It is easy to see that neither the red nor the blue set in Fig. 1 admits a bad rectangle. Therefore the chosen values of the red and blue functions force the remaining four types of atoms to occupy positions outside the red and blue areas.

Next, values of the green and sienna functions are chosen so that the green and sienna parts of the master block represent sets whose X-rays agree with these functions. Again, the sets are unique; an argument can be based on bad rectangles, but we prefer the following alternative method.

On the left-most column on which the vertical green function is nonzero (the third column in Fig. 2), its value is $2(p-1)$. Since there are only $2(p-1)$ lattice squares in this column above the red area on which the horizontal green function has a nonzero value, the position on this column of green atoms in any solution is uniquely determined. The same argument applies to all the other columns for which the vertical green function has value $2(p-1)$ (for example, the eleventh column in Fig. 2 or the first column in Fig. 3). Now consider the fourth column, for which the vertical green function has value $2(p-2)$. The $2(p-2)$ green atoms indicated in Fig. 2 are the only possible ones in a solution, since the horizontal green function has values 1 and 2 on rows $(p-1)$ and $2p$, respectively, and these rows already contain these numbers of green atoms. Again, the same applies to all the other columns for which the vertical green function has value $2(p-2)$ (for example, the twelfth column in Fig. 2 or the second column in Fig. 3). Applying this argument inductively we see that the set of green atoms is uniquely determined in any solution. A similar proof shows that the set of sienna atoms in any solution is also uniquely determined by the sienna function values.

We now turn our attention to the indigo and orange atoms in a solution, which can only occupy squares left free by the red, blue, green, and sienna atoms. The values of the horizontal indigo function are chosen to be 0 on row $i(p+1)$, $i=1,\ldots,3t$, and 1 on all other rows meeting the master block; see Figs. 2 and 3. The values of the vertical indigo function on the first and last columns of the master block are $(p-k)$ and k, respectively, and on the intermediate columns the values repeat a pattern of 0 followed by p 1's.

The values of the orange functions are 0 on all rows and columns except those meeting a component block B_i. In the instance $\mathscr{I}$ of k-EXACT COVER, a certain subset of $\mathscr{P}$ consists of elements containing p_i. If there are r_i such elements in this subset, the value of the horizontal orange function is chosen to be 1 on the corresponding r_i rows among the first p rows of B_i. (In the example depicted in Fig. 2, we have $r_1=4$ and the first, third, fourth, and fifth elements of the seven in $\mathscr{P}$ contain p_1.) The remaining $(p-r_i)$ values of the horizontal orange function on the first p rows are chosen to be 0. The same pattern is repeated for rows $(2p+3),\ldots,(3p+2)$. Row $(p+1)$ is assigned value 1 and row $(3p+3)$ the value (r_i-1). All other rows meeting B_i are given the value 0. The vertical orange function values on columns meeting B_i

conform to a similar pattern of 0's and 1's, with two special columns, the first and $(p+2)$th, having values 1 and (r_i-1), respectively. We refer the reader again to Fig. 2.

This completes the description of the instance $\mathscr{J}$ of POLY$_6$-CONSISTENCY$_{\mathscr{F}^2}(S_1^*, S_2^*)$. Suppose there is a solution, that is, disjoint sets of atoms of the six types whose horizontal and vertical X-rays have the function values chosen above.

The red, blue, green, and sienna atoms in the solution must occupy the positions shown in Figs. 1 and 2, for reasons given above. Observe next that there must be $(p-k)$ indigo atoms in the first column of the master block, and the horizontal indigo function values imply that there are k indigo atoms lying in complementary positions on the diagonal segment D_1 of length p at the top left of B_1 (see Fig. 2). Now the values of the vertical indigo function imply that there are $(p-k)$ indigo atoms lying in the same relative positions as the original ones on another diagonal segment below D_1. This process continues through B_1 and via the first connecting block to B_2, and so on in this fashion throughout the master block. The relative positions of the $(p-k)$ (or k) indigo atoms on each alternate diagonal segment is always the same. Also, the output at the bottom right of the master block contains k indigo atoms. We shall show below that the position of the k indigo atoms on any alternate diagonal segment yield the required solution for the instance $\mathscr{I}$ of k-EXACT COVER.

To prove this we must consider the orange atoms in the solution for $\mathscr{J}$. Recall that the orange function values encode those elements in $\mathscr{P}$ that contain a particular element of P. We must show that in each B_i, exactly one of the k indigo atoms on the diagonal segment at its top left represents an element in $\mathscr{Q}$ that contains p_i; in other words, exactly one of these k indigo atoms lies on a row on which the value of the horizontal orange function is 1. In the example depicted in Fig. 2, this is the one on the fifth row.

Consider any component block, say the first. We have seen that the positions of the red, blue, green, and sienna atoms are fixed. The values of the horizontal orange function imply that there are either (r_1-1) or r_1 orange atoms on D_1. Suppose that there are r_1 such atoms. Considering alternately the values of the vertical and horizontal indigo functions, we see that there are r_1 indigo atoms on the diagonal segment D_2 meeting rows $(2p+3)$ through $(3p+2)$ and columns $(p+3)$ through $(2p+2)$. Furthermore, these indigo atoms must occupy the only permissible lattice squares available on D_2 for orange atoms. However, if there are no orange atoms on D_2, there must be r_1 orange atoms in column $(p+2)$, contradicting the value (r_1-1) of the vertical orange function on this column.

We conclude that there are exactly (r_1-1) orange atoms on D_1. Suppose that there are no indigo atoms on D_1 also lying on a row on which the horizontal orange function has value 1. The indigo function values then imply that there are indigo atoms on D_2 on every column on which the vertical orange function has value 1. This in turn means that there are no orange atoms on D_2, which we saw above is impossible. Therefore, since there are r_1 rows among the first p on which the horizontal orange function has value 1, there must be precisely one such row on which there is an indigo atom

also on D_1. Consequently, we have obtained the desired solution for the instance $\mathscr{I}$ of k-EXACT COVER.

Now suppose we are given a solution $\mathscr{Q}$ for the instance $\mathscr{I}$ of k-EXACT COVER. The red, blue, green, and sienna atoms in the positions shown form a partial solution for the instance $\mathscr{J}$ of POLY$_6$-CONSISTENCY$_{\mathscr{F}^2}(S_1^*, S_2^*)$, since the corresponding X-rays have values equal to those prescribed above. From the set $\mathscr{Q}$ we can determine the positions of all the indigo atoms via the correspondence explained above in such a way that they are also part of a solution for $\mathscr{J}$. For $i = 1,\ldots,t$, there is a unique element in $\mathscr{Q}$ containing p_i. This determines the position of the orange atoms in B_i, again via the above correspondence, so that they too are part of a solution for $\mathscr{J}$. Therefore we have a solution for the instance $\mathscr{J}$ of POLY$_6$-CONSISTENCY$_{\mathscr{F}^2}(S_1^*, S_2^*)$.

Finally, note that since we can assume without loss of generality that $k \leqslant p$, the size of the master block is bounded by a polynomial in p and t. Therefore the above transformation runs in strongly polynomial time. Clearly, it is also parsimonious. □

Putting together Lemma 3.1, Theorem 3.2, the fact that #(k-EXACT COVER) is #$\mathbb{P}$-complete, and the fact that the transformation constructed in the previous proof is parsimonious, we obtain the following result.

Theorem 3.3. *If $c \geqslant 6$, then* POLY$_c$-CONSISTENCY$_{\mathscr{F}^d}(S_1, S_2)$ *is* $\mathbb{NP}$*-complete* (*in the strong sense*), *and its counting version* #(POLY$_c$-CONSISTENCY$_{\mathscr{F}^d}(S_1, S_2)$) *is* #$\mathbb{P}$*-complete.*

It follows immediately that POLY$_c$-RECONSTRUCTION$_{\mathscr{F}^d}(S_1, S_2)$ is $\mathbb{NP}$-hard for $c \geqslant 6$.

We now turn to POLY$_c$-UNIQUENESS$_{\mathscr{F}^2}(S_1, S_2)$. We find it convenient to employ the $\mathbb{NP}$-completeness of the following problem.

UNIQUE SATISFIABILITY

Instance: *Positive integers s, t, a set V of s Boolean variables, a set $\mathscr{C}$ of t clauses over V, and a truth assignment T for $\mathscr{C}$.*

Question: *Is there a satisfying truth assignment for $\mathscr{C}$ that is different from T?*

We believe the follow lemma is known but did not find a proof in the literature. (See [1, 12] for complexity classes and results related to a variant of the above problem: given a Boolean formula, decide whether it has exactly one satisfying truth assignment.)

Lemma 3.4. UNIQUE SATISFIABILITY *is* $\mathbb{NP}$*-complete in the strong sense.*

Proof. Clearly, UNIQUE SATISFIABILITY is in $\mathbb{NP}$. To prove $\mathbb{NP}$-completeness we describe a transformation from SATISFIABILITY.

Suppose $\mathscr{I} = (s, t, V, \mathscr{C})$ is an instance of SATISFIABILITY. We assume that $V = \{x_1,\ldots,x_s\}$ and $\mathscr{C} = \{C_1,\ldots,C_t\}$, so the problem is to decide whether the Boolean expression

$$\mathscr{B} = C_1 \wedge C_2 \wedge \ldots \wedge C_t$$

has a satisfying truth assignment. We shall construct an instance $\mathscr{I}' = (s', t', V', \mathscr{C}', T')$ of UNIQUE SATISFIABILITY such that there is a solution for the former if and only if there is solution for the latter.

Suppose that $C_1 = (l_1 \vee \ldots \vee l_n)$. We may assume that $n \leqslant s$ and each variable occurs only once, so by relabeling the variables, if necessary, we have $l_i \in \{x_i, \neg x_i\}$ for $i = 1, \ldots, n$.

We define an instance $\mathscr{I}' = (s', t', V', \mathscr{C}', T')$ of UNIQUE SATISFIABILITY by letting $s' = s$, $t' = st$, $V' = V$,

$$\mathscr{C}' = \{(C \vee \neg l_i) : C \in \mathscr{C}, i = 1, \ldots, n\} \cup \{(C \vee x_i) : C \in \mathcal{C}, i = n+1, \ldots, s\},$$

and letting T' be the truth assignment for which l_i is false, $i = 1, \ldots, n$, and x_i is true, $i = n+1, \ldots, s$. Note that T' is indeed a satisfying truth assignment for the corresponding Boolean expression

$$\mathscr{B}' = \left(\bigwedge_{i=1}^{t}\bigwedge_{j=1}^{n}(C_i \vee \neg l_j)\right) \wedge \left(\bigwedge_{i=1}^{t}\bigwedge_{j=n+1}^{s}(C_i \vee x_j)\right),$$

so $\mathscr{I}'$ is a proper instance of UNIQUE SATISFIABILITY.

Suppose that there is a solution for $\mathscr{I}$, that is, a satisfying truth assignment T. Since at least one of the literals of C_1 is true in T, we have $T \neq T'$. Moreover, rewriting $\mathscr{B}'$ as

$$\mathscr{B}' = \left(\bigwedge_{i=1}^{t} C_i\right) \vee \left(\bigwedge_{j=1}^{n} \neg l_j \wedge \bigwedge_{j=n+1}^{s} x_j\right),$$

we see that T satisfies $\mathscr{B}'$. Therefore there is a solution for $\mathscr{I}'$.

Conversely, a solution for $\mathscr{I}'$ yields a truth assignment T different from T' that satisfies $\mathscr{B}'$. In T, the expression $\left(\bigwedge_{j=1}^{n} \neg l_j \wedge \bigwedge_{j=n+1}^{s} x_j\right)$ is false, since $T \neq T'$. Therefore $\mathscr{B} = \left(\bigwedge_{i=1}^{t} C_i\right)$ must be true in T, so there is a solution for $\mathscr{I}$.

Finally, note that the transformation runs in strongly polynomial time. □

Theorem 3.5. POLY$_c$-UNIQUENESS$_{\mathscr{F}^d}(S_1, S_2)$ *is* $\mathbb{NP}$*-complete* (*in the strong sense*) *for* $c \geqslant 6$.

Proof. By Lemma 3.1, it suffices to prove the result for POLY$_6$-UNIQUENESS$_{\mathscr{F}^2}(S_1^*, S_2^*)$.

In [5, 15] it is shown that there is a parsimonious transformation from SATISFIABILITY to k-EXACT COVER. By Lemma 3.4, this implies that what may be called UNIQUE k-EXACT COVER is $\mathbb{NP}$-complete in the strong sense. Since the transformation from k-EXACT COVER to POLY$_6$-UNIQUENESS$_{\mathscr{F}^2}(S_1^*, S_2^*)$ constructed in Theorem 3.2 runs in strongly polynomial time and is parsimonious, the result follows. □

4. Remarks and Consequences

4.1. A conjecture

By the results of Theorem 3.3 and the consequences of [6] stated before, the only cases for which the computational complexity of POLY$_c$-CONSISTENCY$_{\mathscr{F}^2}(S_1^*, S_2^*)$ and POLY$_c$-UNIQUENESS$_{\mathscr{F}^2}(S_1^*, S_2^*)$ are left undetermined are those for which $c = 2, 3, 4, 5$.

Conjecture 4.1. POLY$_c$-CONSISTENCY$_{\mathscr{F}^2}(S_1^*, S_2^*)$ *and* POLY$_c$-UNIQUENESS$_{\mathscr{F}^2}(S_1^*, S_2^*)$ *are* $\mathbb{NP}$*-complete* (*in the strong sense*) *for* $c \geqslant 3$.

We believe that a substantially new technique will be needed, at least for the case $c = 3$.

Question 4.2. What is the computational complexity of POLY$_2$-CONSISTENCY$_{\mathscr{F}^2}(S_1^*, S_2^*)$ and POLY$_2$-UNIQUENESS$_{\mathscr{F}^2}(S_1^*, S_2^*)$?

We have not yet seen unpublished work of Picouleau, which may bear on the previous question.

4.2. Separate versus joint uniqueness

We noted above that it is known that when $c = 1$ and $m = 2$, uniqueness can be checked in polynomial time. Under the assumption that $\mathbb{P} \neq \mathbb{NP}$, therefore, Theorem 3.5 shows that when $m = 2$, the problem of checking uniqueness for $c \geqslant 6$ types of atoms is not equivalent to that of checking uniqueness for each type of atom separately.

It remains possible (and relevant to the second part of Question 4.2) that checking uniqueness for two types of atoms is equivalent to checking uniqueness for three related homogeneous problems: the two obtained by considering the data for each type of atom separately and the problem obtained by adding the corresponding functions for the two types of atoms. However, an example from [3], depicted in Fig. 4, shows that even consistency does not follow from the consistency and uniqueness of the three related homogeneous problems. The left and middle parts of Fig. 4 give the unique configurations of red and blue atoms, respectively, having the row and column sums indicated. The right part does the same when these margin sums are added. It is easy to check that there is no solution for the instance in which the red and blue functions are considered together.

Suppose that the data for two types of atoms determine the position of atoms of one type. Then uniqueness can be checked in polynomial time, since we can delete the determined atoms from the candidate set and then check the other type for uniqueness in polynomial time. Now suppose that when the two types of data are summed, this summed data determines uniquely a solution set F. Then we can check in polynomial time within F for a solution for the data for one type of atom. If there is no such solution, there cannot be a solution for the original data, while if there is, the

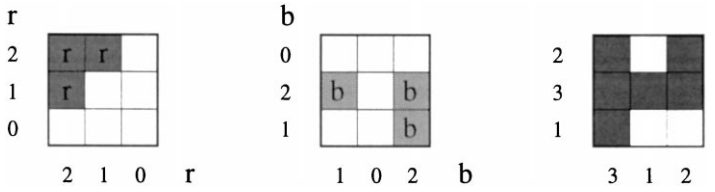

Fig. 4.

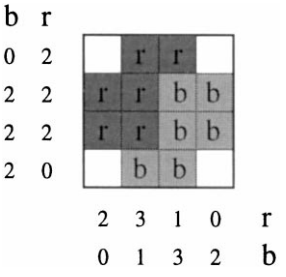

Fig. 5.

remaining type of atom must fill the remaining spaces in F. Generally, however, the following example, together with the previous one, indicates that the relation between separate and joint uniqueness might be rather complicated. Fig. 5 shows an instance of POLY$_2$-CONSISTENCY$_{\mathscr{F}^2}(S_1^*, S_2^*)$ where there are three different solutions for each color separately, yet the specified instance has a unique solution.

4.3. Contingency tables

Theorem 3.2 shows that the consistency problem for the reconstruction of a restricted type of *contingency table* from its row and column sums is $\mathbb{NP}$-complete.

In fact, let q be a prime number larger than the number of rows and the number of columns of the master block in the construction from Theorem 3.2. Using the lattice squares in the master block, we can define an instance $\mathscr{K}$ of a contingency table consistency problem in which the nonzero table entries are restricted to integers of the form q^l, where $l = 0, 1, \ldots, 5$. Let the functions defined above specifying row sums for the six colors be $f_0, f_1, \ldots, f_5$. Then we define a function g_1 by

$$g_1(T) = f_0(T) + f_1(T)q + \cdots + f_5(T)q^5,$$

for each row T. It is easy to see that the value of $g_1(T)$ uniquely determines the numbers $f_l(T)$ in the previous equation. We define a function g_2 on columns analogously. If there is a solution for $\mathscr{K}$, then there is a contingency table with nonzero entries of the form q^l, where $l = 0, 1, \ldots, 5$, whose row sums and column sums are given by g_1 and g_2. A solution for $\mathscr{J}$ is then obtained by identifying each entry q^l with an atom of the corresponding color. Conversely, if there is a solution for $\mathscr{J}$, one obtains

a solution for $\mathscr{K}$ by replacing each atom of a given color by the entry q^l corresponding to that color. We conclude that the consistency problem for this restricted type of contingency table is $\mathbb{NP}$-complete and the corresponding reconstruction problem is $\mathbb{NP}$-hard. Similarly, the uniqueness problem is also $\mathbb{NP}$-complete.

4.4. Data security

The results of the previous section also complement work in [9] concerning three-dimensional statistical *data security*. Here the problem is to reconstruct a three-dimensional table whose entries are nonnegative integers from its row, column, and file sums, that is, the line sums parallel to the x-, y-, and z-axes. In [9] it is shown that this reconstruction problem for $n \times n \times n$ tables is $\mathbb{NP}$-hard even if the line sums are restricted to $\{0,1\}$. (Note that in the context of data security, $\mathbb{NP}$-hardness is highly desirable, of course.)

It follows from Theorem 3.2 that the reconstruction problem is also $\mathbb{NP}$-hard for $n \times n \times 7$ tables when the row and column sums are unrestricted and the file sums are 1 on the corresponding $n \times n$ grid. To see this, consider again the instance $\mathscr{J}$ of POLY$_6$-CONSISTENCY$_{\mathscr{F}^2}(S_1^*, S_2^*)$ constructed in Theorem 3.2, and suppose the size of the master block used there is $m \times n$. We construct an instance $\mathscr{K}'$ of the table reconstruction problem as follows. Each of the first six horizontal levels of the table is assigned one of the six colors used in the proof of Theorem 3.2. The row and column sums for each such level are the same as the horizontal and vertical function values, respectively, specified for that particular color in the instance $\mathscr{J}$. If T is any row in the seventh level of the table, the row sum for T is n minus the sum of the row sums for the rows below T in the first six levels of the table; in other words, the number of lattice squares unoccupied by atoms of any color in any solution for $\mathscr{J}$. Column sums in the seventh level are defined analogously. Finally, the file sum for each file is 1. A solution for $\mathscr{J}$ immediately yields a solution for $\mathscr{K}'$ when 0's and 1's are inserted in the table according to the position of atoms, via the correspondence just described. Conversely, a solution for $\mathscr{K}'$ must have entries that are 0 or 1, in view of the file sums, and the position of the 1's in each of the first six levels indicates the position of the atoms in a solution for $\mathscr{J}$. (Note that the file sums guarantee that at most one atom can occupy any lattice square.) It is now easy to modify the construction to extend the result to the case $m = n$.

4.5. Scheduling

Our reconstruction problem can be interpreted as a special machine scheduling problem. Here each task is composed of several types of units and each machine must perform specified quantities of the various types of units. The problem can be recast in the form POLY$_c$-RECONSTRUCTION$_{\mathscr{F}^2}(S_1^*, S_2^*)$, where c is the number of types of units, now regarded as types of atoms. The machines are identified with a finite subset of $\mathscr{A}(S_1^*)$ and the tasks with a finite subset of $\mathscr{A}(S_2^*)$. The functions $f_{1,l}$ and $f_{2,l}$

encode the units to be performed by each machine and the units comprising each task, respectively. Theorem 3.2 is then applicable.

References

[1] A. Blass, Y. Gurevich, On the unique satisfiability problem, Inform. and Control 55 (1982) 80–88.
[2] S.-K. Chang, The reconstruction of binary patters from their projections, Comm. ACM 14 (1971)
[3] Y.-C. Chen, A. Shastri, On joint realization of (0,1) matrices, Linear Algebra Appl. 112 (1989) 75–85.
[4] P.C. Fishburn, J.C. Lagarias, J.A. Reeds, L.A. Shepp, Sets uniquely determined by projections on axes. II. Discrete case, Discrete Math. 91 (1991) 149–159.
[5] Z. Galil, On some direct encodings of nondeterministic touring machines operating in polynomial time into P-complete problems, SIGACT News 6 (1) (1974) 19–24.
[6] R.J. Gardner, P. Gritzmann, D. Prangenberg, On the computational complexity of reconstructing lattice sets from their X-rays, preprint.
[7] R.J. Gardner, P. Gritzmann, D. Prangenberg, On the reconstruction of binary images from their discrete Radon transforms, in: R.A. Melter, A.Y. Wu, L. Latecki (eds.), Vision Geometry V, Society of Photo-Optical Instrumentation Engineers Proceedings, vol. 2826, 1996, pp. 121–132.
[8] J.J. Gerbrands, C.H. Slump, A network flow approach to reconstruction of the left ventricle from two projections, Comput. Graphics Image Process. 18 (1982) 18–36.
[9] R.W. Irving, M.R. Jerrum, Three-dimensional statistical data security problems, SIAM J. Comput. 23 (1994) 170–184.
[10] R. Karp, Reducibility among combinational problems, in: R.E. Miller, J.W. Thatcher (eds.), Plenum Press, New York, 1972, pp. 85–103.
[11] C. Kisielowski, P. Schwander, F.H. Baumann, M. Seibt, Y. Kim, A. Ourmazd, An approach to quantitative high-resolution transmission electron microscopy of crystalline materials, Ultramicroscopy 58 (1995) 131–155
[12] C.H. Papadimitriou, M. Yannakakis, The complexity of facets (and some facets of complexity), J. Comput. System. Sci. 28 (1984) 244–259.
[13] H.J. Ryser, Combinatorial Mathematics, Mathematical Association of America and Quinn & Boden, Rahway, NJ, 1963.
[14] P. Schwander, C. Kisielowski, M. Seibt, F.H, Baumann, Y. Kim, A. Ourmazd, Mapping projected potential interfacial roughness, and composition in general crystalline solids by quantitative transmission electron microscopy, Phys. Rev. Lett. 71 (1993) 4150–4153.
[15] J. Simon, On some central problems in computational complexity, Ph.D. thesis, Dept. of Computer Science, Cornell University, Ithaca, NY, 1975.